BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 3: e37013
doi: 10.3897/biss.3.37013

OPEN ACCESS

Conference Abstract

Uniquely Identifying Collectors of Specimens

Quentin Groom‡, Mathias Dillen‡, Pieter Huybrechts‡

‡ Meise Botanic Garden, Meise, Belgium

Corresponding author: Quentin Groom (quentin.groom1@plantentuinmeise.be)

Received: 09 Jun 2019 | Published: 13 Jun 2019

Citation: Groom Q, Dillen M, Huybrechts P (2019) Uniquely Identifying Collectors of Specimens. Biodiversity Information Science and Standards 3: e37013. https://doi.org/10.3897/biss.3.37013

Abstract

Being able to connect together the specimens collected and determined by a person is key to understanding the origins, history, identity and ownership of a collection. Yet making these connections is difficult because people's names change, are not unique and are written in many different ways. Semantic annotation of people names with stable identifiers is a solution to this problem.

Semantic annotation requires the linking of textual information to unique identities for known entities. This often needs to be done manually and can be time-consuming and error-prone. We have been linking person text strings to persistent identifiers of various kinds. We focused on the semantic annotation of people names mentioned on specimen labels, particularly the collectors, but also determiners of the specimen. Various person identifiers systems exist and they vary in their scope, their business model and their notability requirements. We have chosen, where possible, to link to several systems, because we feel this creates a more robust system, that does not rely on the maintenance of any single identifier system.

Currently, we are using Virtual International Authority File (VIAF) identifiers that are particularly useful for published authors. The Harvard University Index of Botanists has information on well known botanists and ORCID IDs are useful for living collectors. However, even with all these various identifiers systems, there are many collectors of whom we know little, particularly where they have not published or done something notable other than collect specimens. For many of these collectors Wikidata is a solution, because it can be edited by anyone and because it has a low threshold for notability. Wikidata also

© Groom Q et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

has the additional advantage of forming a bridge between many identifier systems. Hence, we incorporated Wikidata entries as well if available and endeavoured to merge existing duplicates wherever we could. In this process, we have examined the coverage, scope and reliability of these different systems for identifying collectors and determiners. We also advise on strategies for linking collectors to their identifiers efficiently. Linking specimens to clearly identified people has enormous advantages for specimen data curation and linking between different collections. We encourage every collection to prioritize this.

Keywords

botanists, ORCID, Virtual International Authority File, VIAF, Wikidata

Presenting author

Quentin Groom

Presented at

Biodiversity_Next 2019

Funding program

H2020 project Synthesys+ under grant agreement 823827